NPS-OR-96-009

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Waiting Times When Service Times are Stable Laws: Tamed and Wild

by

Donald P. Gaver
Patricia A. Jacobs

August 1996

Approved for public release; distribution is unlimited.

Prepared for: Naval Postgraduate School
Monterey, CA 93943-5000

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943-5000

Rear Admiral M. J. Evans
Superintendent

Richard Elster
Provost

This report was prepared for and funded by the Naval Postgraduate School.

Reproduction of all or part of this report is authorized.

This report was prepared by:

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE August 1996	3. REPORT TYPE AND DATES COVERED Technical

4. TITLE AND SUBTITLE Waiting Times When Service Times are Stable Laws: Tamed and Wild	5. FUNDING NUMBERS
6. AUTHOR(S) Donald P. Gaver and Patricia A. Jacobs	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943	8. PERFORMING ORGANIZATION REPORT NUMBER NPS-OR-96-009
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	10. SPONSORING / MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES	
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.	12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Modern telecommunication systems must accommodate tasks or messages of extremely variable time duration. Understanding of that variability, and appropriate stochastic models are needed to describe the resulting queues or buffer contents. To this end, consider an M/G/1 queue with service times having a positive stable law distribution. Such service times are extremely long (and short) tailed, and thus do not have finite first and second moments; classical queue-theoretic results do not apply directly. Here we suggest two procedures for initially *taming* stable laws, i.e. so that they possess finite mean and variance. We apply the tamed laws to calculate certain familiar queuing properties, such as the transform of the stationary distribution of the long-run virtual waiting time and mean thereof. We show that, by norming or scaling traffic intensity, waiting times, and other measures of congestion, we can obtain *bona fide* limiting distributions as the underlying service times become untamed, i.e. return to the wild. Simulations support the theory.

14. SUBJECT TERMS queuing, positive stable law service times, virtual waiting time, limiting results	15. NUMBER OF PAGES 30
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL

Waiting Times When Service Times are Stable Laws: Tamed and Wild

Donald P. Gaver
Patricia A. Jacobs

August 1996

To appear in ***Recent Contributions in Applied Probability and Stochastic Process,*** (A book in honor of Julian Keilson).

WAITING TIMES WHEN SERVICE TIMES ARE STABLE LAWS: TAMED AND WILD

Donald P. Gaver

Patricia A. Jacobs

Department of Operations Research
Naval Postgraduate School
Monterey, CA 93943

Abstract

Modern telecommunication systems must accommodate tasks or messages of extremely variable time duration. Understanding of that variability, and appropriate stochastic models are needed to describe the resulting queues or buffer contents. To this end, consider an M/G/1 queue with service times having a positive stable law distribution. Such service times are extremely long (and short) tailed, and thus do not have finite first and second moments; classical queue-theoretic results do not apply directly. Here we suggest two procedures for initially *taming* stable laws, i.e. so that they possess finite mean and variance. We apply the tamed laws to calculate certain familiar queuing properties, such as the transform of the stationary distribution of the long-run virtual waiting time and mean thereof. We show that, by norming or scaling traffic intensity, waiting times, and other measures of congestion, we can obtain *bona fide* limiting distributions as the underlying service times become untamed, i.e. return to the wild. Simulations support the theory.

1. The Problem Motivation

In various applications of service system or queuing theory there may arise a need to consider service times, S, of great variability, i.e. that seem to possess nearly Pareto tails:

$$P\{S > x\} \equiv 1 - F_S(x) = O\left(x^{-\alpha}\right) \tag{1.1}$$

as $x \to \infty$, where α is small enough so that no moments, $E[S^k]$, $k \geq 1$, are finite. In this paper we examine certain aspects of such problems for M/G/1 systems, focusing on service times that are describable by *positive stable laws.* In view of Theorem 1 on p. 448, Feller, II (1971), it is impossible to ignore the class of stable law models to represent the behavior of (1.1); there is the additional fact that stable laws approximate the distributions of sums of many long-tailed independent random variables, e.g. the sum of a number of activities that constitute service. But there is the problem that without finite first and second moments at a minimum classical queue-theoretic results do not directly apply.

In this paper we consider some procedures for *taming* stable laws so that they do possess the required properties, i.e. finite moments. We apply the tamed laws to calculate certain familiar queuing properties, such as the virtual waiting time in the system. Then we show that, by norming or scaling waiting times and other measures of congestion, we can obtain bona-fide limiting distributions as the underlying service times become untamed, or "return to the wild". For similar work see Abate, *et al.* (1993, 1994) and probably more recent articles as well.

The authors are very much indebted to Walter Willinger for pointing out many interesting references attesting to the appearance of long-tailed distributions in modern communications systems. This in no way implicates W. Willinger in our present machinations. We also gratefully acknowledge the work by Ward Whitt and Joe Abate.

2. How to Tame a Wild Stable Law

There are several approaches that naturally suggest themselves for endowing a stable-law distributed service time, S, of scale parameter, ν, and order $0 \leq \alpha < 1$, with finite moments. Recall from Feller (1971) that the Laplace-Stieltjes transform of S is

$$E\left[e^{-sS}\right] = \exp\left[-(\nu s)^{\alpha}\right]. \tag{2.1}$$

For the value $\alpha = 1/2$ the above possesses an explicit inverse, the name of which is *inverse Gaussian,* a slight misnomer since the distribution is actually that of the inverse square root of a Gaussian. Otherwise inverses are only expressible as unintelligible infinite series or the equivalent.

(2.1) Assessing Shape Indirectly

As stated, (wild) stable laws possess no finite moments of order ≥ 1. Furthermore, there are no conveniently obtained explicit quantiles (e.g. median, lower or upper quantiles, etc.) of the above, but there are simple substitutes based on exponential distributions: ask for the *test* or *killing* exponential density, of mean $\kappa(p)$, an observation from which, X, exceeds S with probability p. We get from (2.1)

$$P\{S < X\} = \exp\left[-\left(\nu/\kappa(p)\right)^{\alpha}\right] = p \tag{2.2}$$

or

$$\kappa(p) = \nu\left[\ln(1/p)\right]^{-1/\alpha}.$$

It is seen that the *exponential median,* $\kappa(1/2)$, approaches ∞ as $\alpha \to 0$, and approaches $\nu/\ln(2)$ as $\alpha \to 1$, not surprisingly since for $\alpha = 1$ the value of $S = \nu$ with probability 1. The value $p = e^{-1} = 0.368$ is pivotal: for $p = e^{-1}$, $\kappa\left(e^{-1}\right) = \nu$ for all α; for $p < e^{-1}$ $(1/p > e)$ $\kappa(p)$ increases with $\alpha \uparrow 1$; for $p > e^{-1}$ $(1/p < e)$ $\kappa(p)$ decreases as $\alpha \uparrow 1$.

(2.2) I, Taming by Tilting: Initial Screening

Large deviation theory exploits an exponential tail by positive tilting towards large values of interest so that the central limit theorem can be applied. Here it is useful to apply *negative tilting,* see Abate *et al.,* (1995) and (1994); they call this *exponential damping,* while we speak of *taming.* Look at S-values that terminate before being killed:

$$P\{S_\kappa \le t\} = P\{S \le t | S \le X\} = \int_0^t f_S(x) e^{-x/\kappa} dx \Big/ \int_0^\infty f_S(x) e^{-x/\kappa} dx \tag{2.3}$$

which has transform

$$E\left[e^{-sS_\kappa}\right] = \exp\left[-(\nu(s+1/\kappa))^\alpha\right] \Big/ \exp\left[-(\nu/\kappa)^\alpha\right]. \tag{2.4}$$

In a queuing context the above might arise naturally as a control strategy: κ, the mean of the service-killing distribution, is selected so as to keep the sizes of the jobs selected under control. The subset of jobs that pass the exponential killing screen are actually allowed into service, so if λ is the arrival rate then the system only sees $\lambda \cdot e^{-(\nu/\kappa)^\alpha}$ as arrival rate, and (unscaled) traffic intensity is

$$\begin{aligned} \rho(\kappa) &= \lambda P\{S < X\} E[S | S < X] \\ &= \lambda e^{-(\nu/\kappa)^\alpha} \cdot \alpha (\nu/\kappa)^{\alpha-1} \nu \qquad (2.5) \\ &\sim \lambda \alpha \nu^\alpha \kappa^{1-\alpha} \to \infty \end{aligned}$$

as $\kappa \to \infty$. Only if $\lambda \kappa^{1-\alpha} = O(1)$ is there hope of achieving a steady-state distribution.

(2.3) II, Taming by Truncation: On-Line Completion, Perhaps Partial

Suppose that each time an S-value is realized a killing (or interruption) value X is independently realized. Total service is rendered if the service survives, i.e. $S < X$; otherwise partial service $X < S$ is rendered and a new job can be accepted as soon as either event occurs. This setup can be called on-line real-time killed service. For the server, it means that the effective service time is $\tilde{S}_\kappa = \min(S, X)$, with transform

$$E\left[e^{-s\tilde{S}_\kappa}\right] = \int_0^\infty e^{-sx}e^{-x/\kappa}f_S(x)dx + \int_0^\infty e^{-sx}(1-F_S(x))e^{-x/\kappa}\,1/\kappa\,dx$$

$$= e^{-(\nu(s+1/\kappa))^\alpha} + \frac{1-e^{-(\nu(s+1/\kappa))^\alpha}}{1+\kappa s} \tag{2.6}$$

$$= \frac{1}{1+\kappa s} + \frac{\kappa s\,e^{-(\nu(s+1/\kappa))^\alpha}}{1+\kappa s}.$$

Now

$$E\left[\tilde{S}_\kappa\right] = \kappa\left(1-e^{-(\nu/\kappa)^\alpha}\right) \tag{2.7}$$

$$\sim \nu^\alpha\kappa^{1-\alpha} \to \infty$$

as the mean killing time $\kappa \to \infty$, so again only if $\lambda\kappa^{1-\alpha} = O(1)$ will there be an opportunity for long-run queue stability.

3. Transforms of Long-Run Waiting Times

The formula for the Laplace-Stieltjes transform of the long-run or steady-state distribution of M/G/1 virtual waiting time, W, is well known to be

$$E\left[e^{-sW}\right] = \frac{1-\rho}{1-\rho\left(\frac{1-E\left[e^{-sS}\right]}{sE[S]}\right)} \tag{3.1}$$

provided $\rho = \lambda E[S] < 1$; otherwise no such distribution exists and the waiting time tends to increase. Now suppose we contemplate an M/G/1 system with stable law service, tamed as in I or II above, i.e. with exponential, X, screening or truncating exponentials, such that $E[X] = \kappa = 1/\mu$. Then consider a sequence of such, as $\kappa \to \infty$ or $\mu \to 0$. We show how to adjust the arrival rate and normalize the waiting time so as to obtain (transforms of) *bona fide* limiting distributions for the normalized virtual waiting times.

First address the scaling of arrival rate λ to control the traffic intensity ρ. From (2.5) and (2.7) it is necessary that the actual arrival rate becomes small as the taming parameter κ becomes large if the resulting traffic intensity is to be bounded. Therefore take the adjusted arrival rate to be $\lambda^* = \lambda\kappa^{1-\alpha}$ constant; the constant is chosen so that the relevant traffic intensity is less than 1. For the screening situation, *I*,

$$\rho_I^* = \lambda^* e^{-(\nu/\kappa)^\alpha} \alpha\nu^\alpha \sim \lambda^* \alpha\nu^\alpha \tag{3.2}$$

and for the truncation situation, *II*,

$$\rho_{II}^* = \lambda^* \kappa^\alpha \left(1 - e^{-(\nu/\kappa)^\alpha}\right) \sim \lambda^* \nu^\alpha. \tag{3.3}$$

For particular stable law input *I*-taming results in smaller system load than does *II*-taming since $\alpha \le 1$. This is to be expected, as the latter admits some arrivals that the former rejects outright.

Assuming the above, consider the normalized random variable $W^* = W/\kappa \equiv W\mu$. Replace s by $\theta/\kappa \equiv \theta\mu$ to obtain

$$E\left[e^{-\theta W^*}\right] = \frac{1-\rho^*}{1-\rho^*\left(\frac{1-E\left[e^{-\theta\mu S}\right]}{\theta\mu E[S]}\right)} \tag{3.4}$$

where S is tamed and $\rho^* < 1$.

(3.1) Screened Service, I

Substitute (2.4) and the expression $E[S_\kappa] = e^{-(\nu/\kappa)^\alpha}\alpha(\nu/\kappa)^{\alpha-1}\nu \sim \alpha\nu^\alpha\mu^{\alpha-1}$ into (3.4). The result is a formula for every screening level $\kappa = 1/\mu$. Now take the limit as $\mu \to 0$:

$$E\left[e^{-\theta W_I^*}\right] = \frac{1-\rho_I^*}{1-\rho_I^*\left(\frac{(1+\theta)^\alpha - 1}{\alpha\theta}\right)}. \tag{3.5}$$

It is clear from construction and also from directly expanding that

$$\psi(\theta) = \frac{(1+\theta)^{\alpha} - 1}{\alpha\theta} \tag{3.6}$$

is completely monotone, hence the transform of an honest distribution. By differentiation or otherwise

$$E\left[W_I^*\right] = \frac{\rho_I^*}{1-\rho_I^*} \cdot \left(\frac{1-\alpha}{2}\right). \tag{3.7}$$

It is immediately seen that the limiting distribution of the scaled limiting random variable W_I^* does not depend on v, the original stable law scale, except through the traffic intensity $\rho^* = \rho_I^* < 1$.

(3.2) Truncated Service, II

The effect of on-line service truncation is traced by substituting the transform (2.6) into (3.4). Take the limit as $\mu \rightarrow 0$, i.e. untame, to obtain

$$E\left[e^{-\theta W_{II}^*}\right] = \frac{1-\rho_{II}^*}{1-\rho_{II}^*\left(\frac{1}{1+\theta}\right)^{1-\alpha}} \tag{3.8}$$

for $\rho_{II}^* < 1$. This is recognized to be the transform of a geometric mixture of gammas with scale 1 and shape parameter $1-\alpha$. In this case

$$E\left[W_{II}^*\right] = \frac{\rho_{II}^*}{1-\rho_{II}^*}(1-\alpha). \tag{3.9}$$

Once again the scaled limiting random variable has a distribution that depends on the service time scale parameter, v, only through the traffic intensity. The fact that $\rho_I^* = \alpha\rho_{II}^* \le \rho_{II}^*$ and that a factor of $1/2$ is present attests to the fact that greater load is placed on system *II* than on system *I*. Of course greater service of all incoming arrivals is furnished by *II* than by *I*.

4. The Number of Customers in the System

The formula for the generating function of the long-run or steady-state distribution of the number of customers waiting or being served at an arbitrary time in an M/G/1 queue, N, is known to be

$$E\left[z^N\right] = \frac{(1-\rho)(1-z)E\left[e^{-\lambda(1-z)S}\right]}{E\left[e^{-\lambda(1-z)S}\right]-z} \tag{4.1}$$

where S is a generic service time and $\rho < 1$; (cf. Gaver [1959]).

4.1 I, Taming by Tilting: Initial Screening

Differentiating the transform of S, (2.4), and evaluating the results at $s = 0$ results in

$$E[S] \sim \alpha \nu^{\alpha} \kappa^{1-\alpha} \tag{4.2}$$

$$Var[S] \sim \alpha(1-\alpha)\nu^{\alpha}\kappa^{2-\alpha} \tag{4.3}$$

as $\kappa \to \infty$. Thus,

$$E[N_I] \sim \frac{\rho_I^{*2}}{2\left(1-\rho_I^{*}\right)} \frac{(1-\alpha)}{\alpha \nu^{\alpha}} \kappa^{\alpha}. \tag{4.4}$$

Substitute (2.4) into (4.1) for $z = e^{-s/\kappa^{\alpha}}$

$$E\left[e^{-sN/\kappa^{\alpha}}\right] \sim \frac{(1-\rho_I)\left(\frac{s}{\kappa^{\alpha}}\right)\exp\left\{-\left[\frac{\nu}{\kappa}\left[1+\lambda^{*}s\right]\right]^{\alpha}+\left(\frac{\nu}{\kappa}\right)^{\alpha}\right\}}{\exp\left\{-\left(\frac{\nu}{\kappa}\left[1+\lambda^{*}s\right]\right)^{\alpha}+\left(\frac{\nu}{\kappa}\right)^{\alpha}\right\}-e^{-s/\kappa^{\alpha}}}$$

$$\to \frac{\left(1-\rho_I^{*}\right)}{1-\rho_I^{*}\left[\frac{\left[1+\lambda^{*}s\right]^{\alpha}-1}{\alpha\lambda^{*}s}\right]} = \frac{\left(1-\rho_I^{*}\right)}{1-\rho_I^{*}\psi\left(\lambda^{*}s\right)} \tag{4.5}$$

as $\kappa \to \infty$, where ψ is defined in (3.6). Note that scaling for N is by κ^{α}, while for W it is by κ. Otherwise (3.5) and (4.5) differ only by a factor λ^{*} in the denominator.

4.2 II, Taming by Truncation

Differentiation of the Laplace transform of the service time (2.6) yields

$$E[S] \sim \nu^{\alpha}\kappa^{1-\alpha} \tag{4.6}$$

$$E[S^2] \sim 2(1-\alpha)\nu^{\alpha}\kappa^{2-\alpha} \tag{4.7}$$

as $\kappa \to \infty$. Thus,

$$E[N_{II}] \sim \frac{\rho_{II}^{*2}}{\left(1-\rho_{II}^{*}\right)} \frac{(1-\alpha)}{\nu^{\alpha}}\kappa^{\alpha}. \tag{4.8}$$

Note that $E[N_I] < E[N_{II}]$ as expected.

Substituting (2.6) into (4.1) for $z = e^{-s/\kappa^{\alpha}}$, it follows that

$$\lim_{\kappa\to\infty} E\left[e^{-sN/\kappa^{\alpha}}\right] = \frac{\left(1-\rho_{II}^{*}\right)}{1-\left(\nu\left[\lambda^{*}s+1\right]\right)^{\alpha}\frac{\lambda^{*}}{1+\lambda^{*}s}}$$

$$= \frac{\left(1-\rho_{II}^{*}\right)}{1-\lambda^{*}\nu^{\alpha}\left[\lambda^{*}s+1\right]^{\alpha-1}} = \frac{\left(1-\rho_{II}^{*}\right)}{1-\rho_{II}^{*}\left(\frac{\left(1/\lambda^{*}\right)}{\left(1/\lambda^{*}\right)+s}\right)^{(1-\alpha)}};$$

this is recognizable as the transform of a geometric mixture of gammas with scale $\frac{1}{\lambda^{*}}$ and shape parameter $1-\alpha$, note its similarity to (3.8). Again the scaling by κ^{α} is involved.

To date inversion of the transform appearing in the denominator of (3.5) and (4.5) has eluded us. We pose the problem of its inversion, or characterization, to Julian Keilson as a birthday gift. Happy Birthday!

5. Busy Periods

It has been seen that normalization by powers of κ, the mean truncation time, permits convergence of the traffic intensity parameter, ρ, and also the stationary distribution of virtual waiting time. It is of interest to study the behavior of the busy period when such a normalization is applied. Here convergence to nice distributions does not occur.

Recall that if B is a busy period duration we can look at its generation in these terms:

$$B = S + B_1 + B_2 + \ldots + B_{N(S)} \tag{5.1}$$

where S is the first service time in the busy period, $\{B_i,\ i = 1, 2, \ldots\}$ is an iid sequence of copies of busy periods starting with one arrival, and $N(S)$ is the number of arrivals in S. By conditional expectation,

$$E\big[B\big|S, N(S)\big] = S + E[B_1]N(S) \tag{5.2}$$

and so

$$E[B] = \frac{E[S]}{1 - \lambda E[S]}. \tag{5.3}$$

If we normalize so that $\lambda = O(1/\kappa^{1-\alpha})$ as in (3.2) and (3.3) then the traffic intensity tends to a constant as κ increases. It follows from (5.3) above that the expected busy period is $E[B]$, like $E[S]$, of order $\kappa^{1-\alpha}$. This gives hope that the actual distribution of a scaled random busy period, $B^{\#} = B/\kappa^{1-\alpha}$, might converge to some recognizable honest form. However, such does not seem to occur. For positive θ, $\varphi_B(\theta)$ is the (smallest positive) root of

$$E\left[e^{-\theta B}\right] \equiv \varphi_B(\theta) = u\left[\theta + \lambda\left(1 - \varphi_B(\theta)\right)\right] \tag{5.4}$$

where $u(\theta)$ is the Laplace-Stieltjes transform of the service time. Hence, the normalized busy period would satisfy, for Model I,

$$\varphi_B^{\#}(\theta) = \exp\left[-\left(\nu\left(\frac{\theta + \lambda^*}{\kappa^{1-\alpha}}\left(1 - \varphi_B^{\#}(\theta)\right) + \frac{1}{\kappa}\right)^{\alpha}\right)\right]\exp\left[(\nu/\kappa)^{\alpha}\right]. \tag{5.5}$$

Differentiation once at $\theta = 0$ shows that

$$E\left[B^{\#}\right] \sim \frac{\alpha v^{\alpha}}{1-\rho_I^*}; \tag{5.6}$$

a similar result, again finite, holds for Model II. However, further analysis shows that for Model I

$$Var\left[B^{\#}\right] \sim \frac{\alpha(1-\alpha)v^{\alpha}\kappa^{\alpha}}{\left(1-\rho_I^*\right)^3} \tag{5.7}$$

i.e. is unbounded as $\kappa \to \infty$, even though it has been normalized and the normalized mean is finite. Similarly, for Model II

$$Var\left[B^{\#}\right] \sim \frac{2(1-\alpha)v^{\alpha}\kappa^{\alpha}}{\left(1-\rho_{II}^*\right)^3} \tag{5.8}$$

which also becomes large like κ^{α} but remains larger than the previous variance because of more permissive job entry. Recall that the traffic intensities in (5.7) and (5.8) differ; refer to (3.2) and (3.3). In summary, it does not appear possible to scale stable-law-service busy periods so as to achieve a non-zero mean and yet get an honest limiting distribution with finite second moment. This is not surprising in light of the fact that the virtual waiting time must be scaled to obtain such a limit. Nevertheless some qualitative information may be deduced about aspects of system behavior from the likes of (5.7) and (5.8).

6. Simulation Results

In this section we describe a simulation experiment and its results.

We consider an M/G/1 queue; the service times have an inverse Gaussian distribution that is tamed by truncation. The transform of the untamed distribution is (2.1) with $\alpha = 1/2$ and $v = 2$. The tamed-by-truncation service time is simulated by

$$S = \min\left(\frac{1}{Z^2}, Y\right)$$

where $\mathbf{Z}$ is a standard normal random variable and Y is an exponential random variable having mean κ.

The customer arrival rate is determined as follows. Set λ^* in (3.3) equal to 0.8. Put the arrival rate of customers

$$\lambda = \lambda^* \Big/ \kappa\left(1 - e^{-(2/\kappa)^{1/2}}\right).$$

The waiting times for successive customers are obtained by recursion

$$W_{n+1} = \max(W_n + S_n - A_{n+1}, 0)$$

where W_n is the n^{th} customer's waiting time in queue, S_n is the length of the n^{th} customer's service time, and A_{n+1} is the time between the n^{th} and $(n + 1)^{\text{st}}$ arrival. Start at $W_1 = 0$. Clearly the above does not simulate *virtual* waiting times, but in the case of Poisson arrivals the long-run limiting results are equivalent.

Graphical displays of the time series of simulated waiting times appear in Figures 1a–1e. One is struck by the large variability in the waiting times: upward surges appear to occur occasionally, prevail for awhile, and then be interrupted by periods of rather small but fluctuating values. Even averages of 15,000 in single realizations are not especially stable: the five quoted range, after normalization by $\kappa = 150$, from 1.48 to 2.46.

Summaries of 5 replications of the simulation appear in Table 1. In each replication the waiting times for 15,000 customers are simulated, and the waiting times for all customers then averaged. These results are reported: the normed-by-$1/\kappa$ averages per replication, when averaged, turn out to equal 2, with standard error of 0.17. This is in excellent agreement with the result of the theory (3.9), which predicts a value of 2.

TABLE 1
LIMITING (κ LARGE) SIMULATION
Mean Waiting Time (Scaled)
15000 Waiting Times per Replication
(TAMING II)
$\kappa = 150$

	1	2	3	4	5	Mean of Means	Standard Error: $\sqrt{Var/5}$
Mean	317.0	369.5	324.1	221.4	269.1	300.2	25.3
$\frac{\text{Mean}}{\kappa}$	2.11	2.46	2.16	1.48	1.79	2.00	0.17

7. Discussion

Taming, as described above, may be viewed as a control strategy. It could be of interest to ask about the fate of those jobs that are rejected (Model I), or partially finished (Model II): these or their residues, respectively, could be shunted to another server that must handle such overflowing extremely long jobs; presumably these occur at a low enough rate to be accommodated because they are filtered from the mainstream of arrivals. Several such stages could be envisioned, and an attempt made to optimize with respect to the taming or truncation parameters κ_s at stages $s = 1, 2, \ldots$. In practice a deterministic truncation time would be realistic, but the mathematics is less tractable.

Finally, we point out that Pareto-tailed distributions are not the most pathologically long-tailed possible. A simple option is to mix one positive stable law with another: replace the parameterization (2.1) by $v^{\#} = v_1^{\alpha}$, where $v^{\#}$ is itself stable. The result is expressible as the Laplace transform of the mixing distribution.

References

Abate, J., Choudhury, G.L., and Whitt, W., "Calculation of the GI/G/1 waiting time distribution and its cumulants from Pollaczek's formulas," *Archiv für Elektronik und Übertragungstechnik* (1993) pp. 311-321.

Abate, J., Choudhury, G.L., and Whitt, W., "Waiting-time tail probabilities in queues with long-tail service-time distributions," *Queueing Systems*, **16** (1994) pp. 311-338.

De Meyer, A. and Teugels, J.L., "On the Asymptotic Behavior of the Distributions of the Busy Period and Service Time in M/G/1," *J. Applied Probab.*, **17** (1980) pp. 802-813.

Duffy, D.E., McIntosh, A.A., Rosenstein, M., and Willinger, W., "Statistical Analysis of CCSN/SS7 Traffic Data from Working CCS Subnetworks," *IEEE Journal on Selected Areas in Communications*, Vol. 12, No. 3, (1994) pp. 544-551.

Erramilli, A., Gordon, J., and Willinger, W., "Applications of Fractals in Engineering for Realistic Traffic Processes," Proceedings of ITC'94, J. Labetoulle and J.W. Roberts (Eds.), Elsevier Science B.V., Amsterdam, The Netherlands (1994) pp. 35-44.

Feller, W., *An Introduction to Probability Theory and Its Applications*, Vol. II, Second Edition, J. Wiley & Sons, Inc. New York (1971).

Gaver, Jr., D.P., "Imbedded Markov chain analysis of a waiting-line process in continuous time," *Ann. Math. Statist.* **30** (1959) pp. 698-720.

Likhanov, N., Tsybakov, B. and Georganas, N.D., "Analysis of an ATM Buffer with Self-Similar ("Fractal") Input Traffic," Proceedings INFOCOM'95, **3** (1995) pp. 985-992.

Pruthi, P. and Erramilli, A., "Heavy-Tailed ON/OFF Source Behavior and Self-Similar Traffic," Proceedings ICC'95 (June 1995) (to appear).

Willekens, E. and Teugels, J.L., "Asymptotic Expansions for Waiting Time Probabilities in an M/G/1 Queue with Long-Tailed Service Time," *Queueing Systems*, **10** (1992) pp. 295-312.

Willinger, W., "Traffic Modeling for High-Speed Networks: Theory versus Practice," invited chapter in Stochastic Networks, F.P. Kelly and R.J. Williams (Eds.), *IMA Volumes in Mathematics and its Applications*, Springer-Verlag, (1995) (to appear).

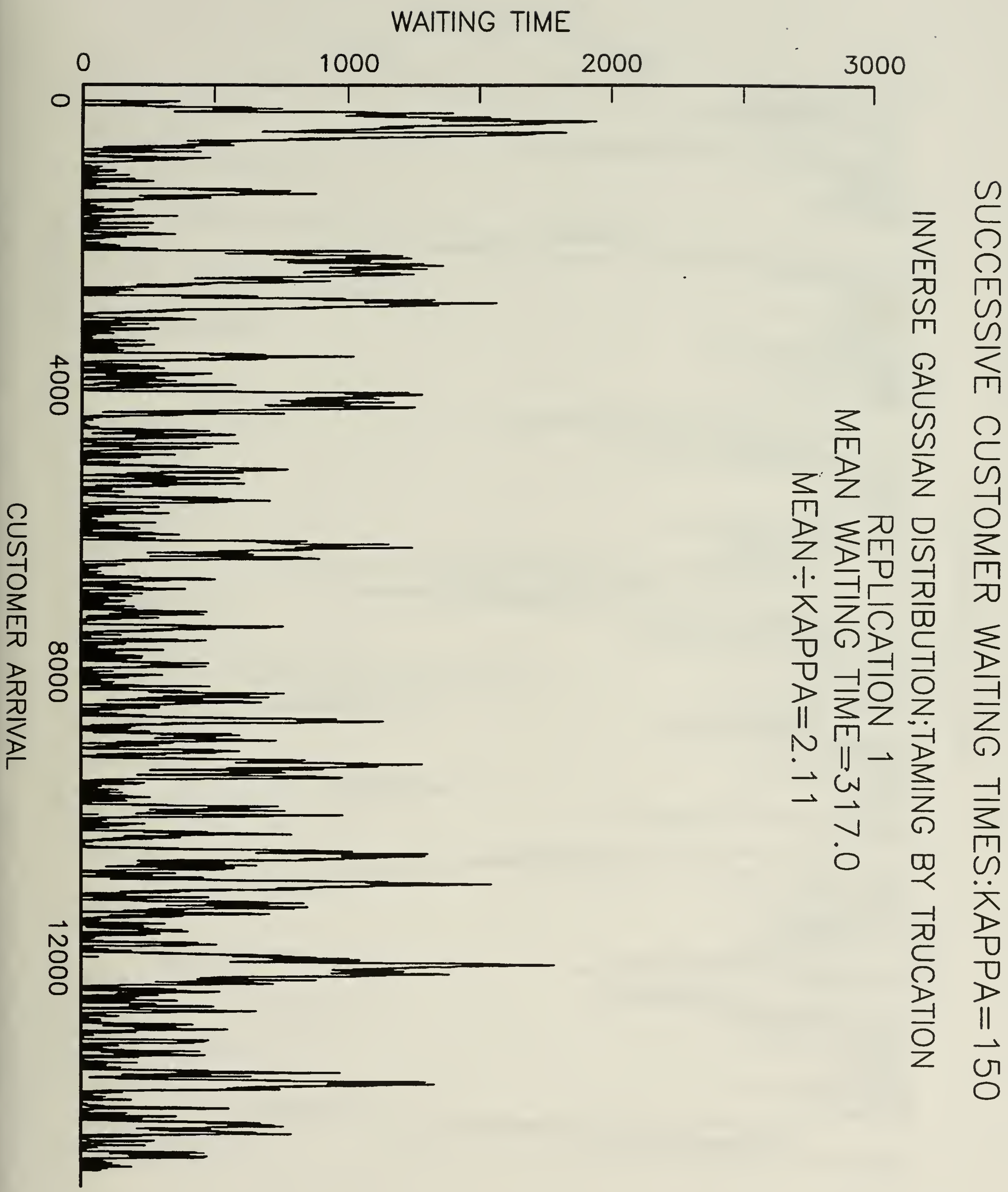

SUCCESSIVE CUSTOMER WAITING TIMES:KAPPA=150
INVERSE GAUSSIAN DISTRIBUTION;TAMING BY TRUCATION
REPLICATION 1
MEAN WAITING TIME=317.0
MEAN÷KAPPA=2.11
WAITING TIME
0
1000
2000
3000
CUSTOMER ARRIVAL
0
4000
8000
12000

Figure 1a

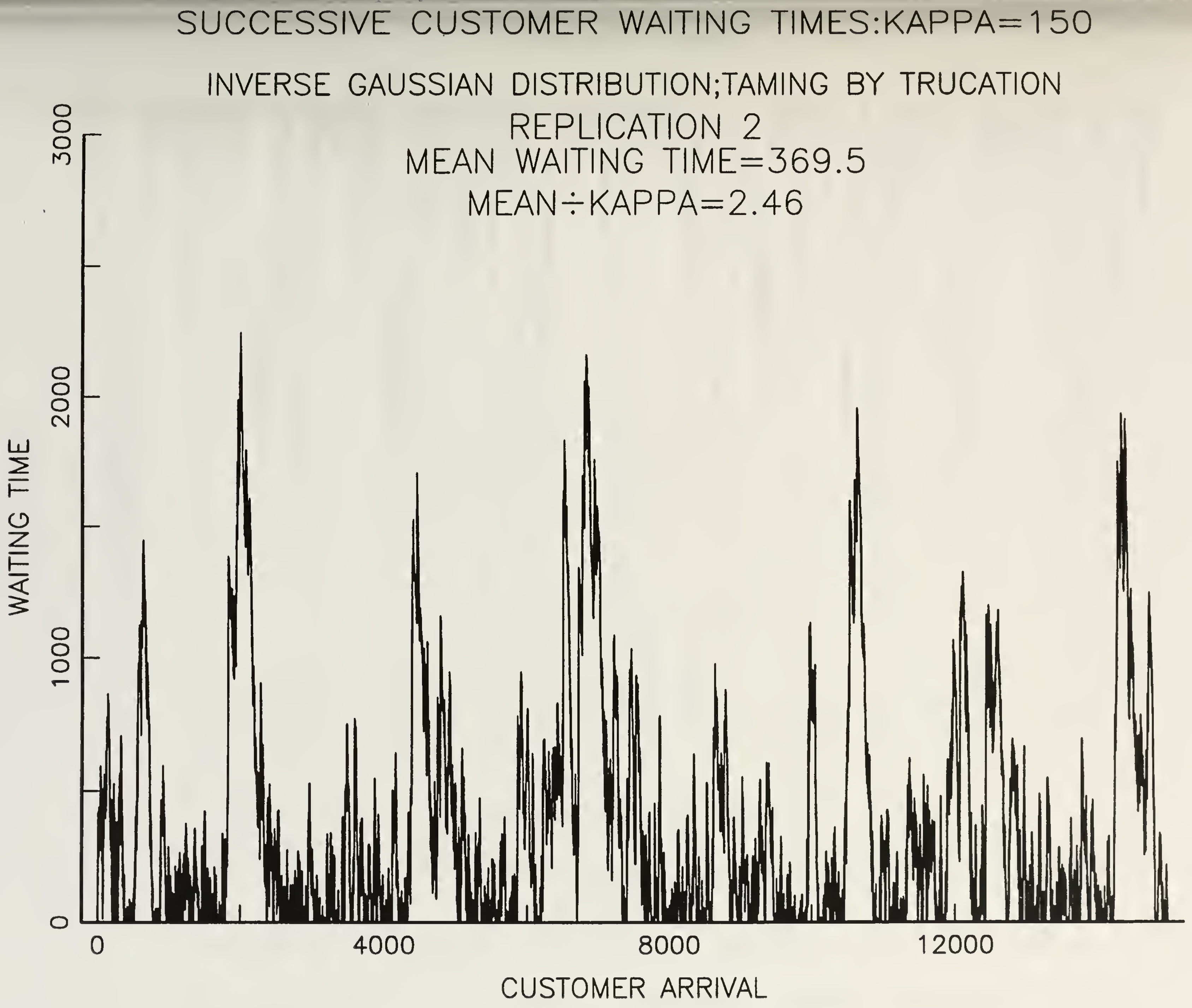
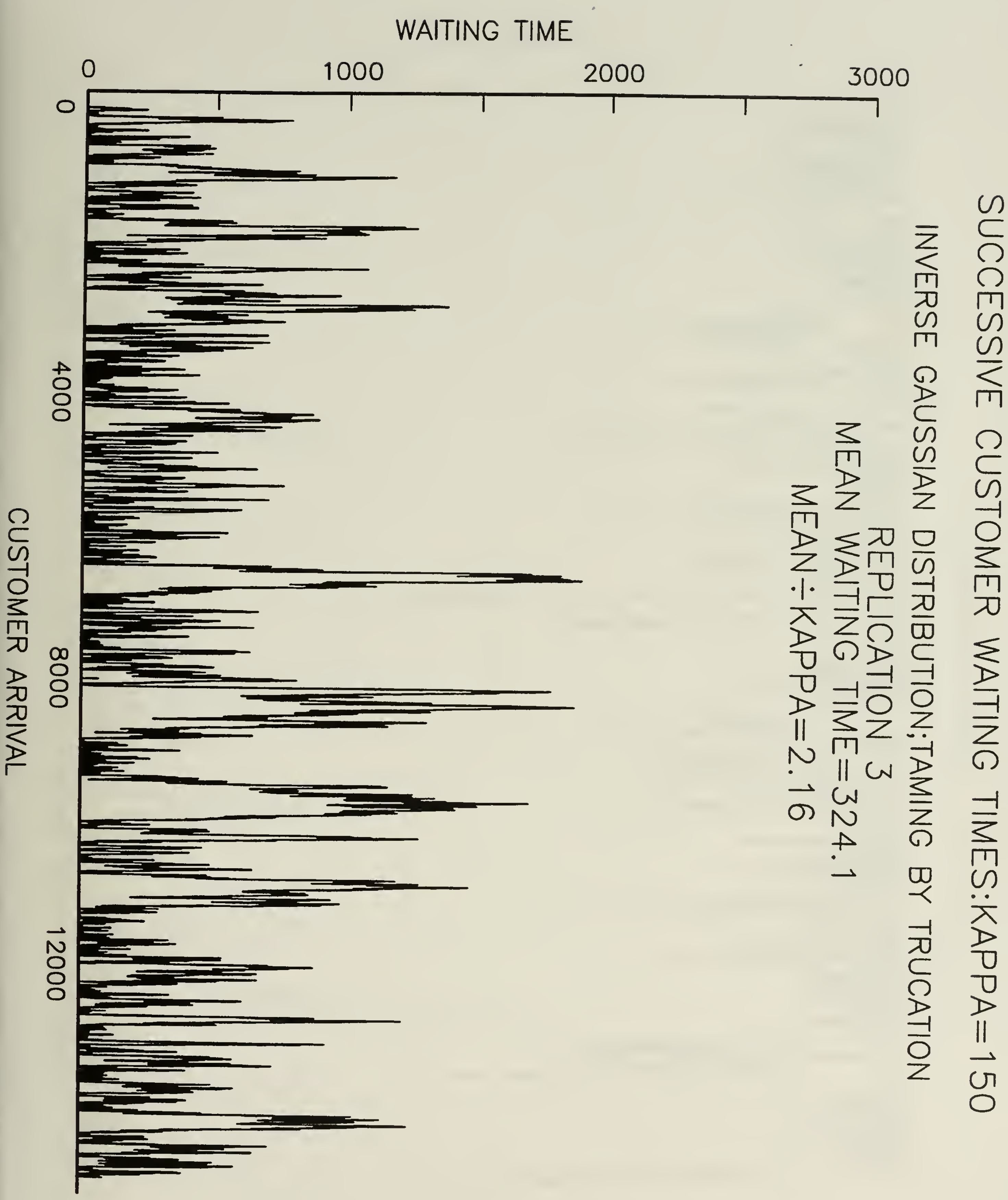
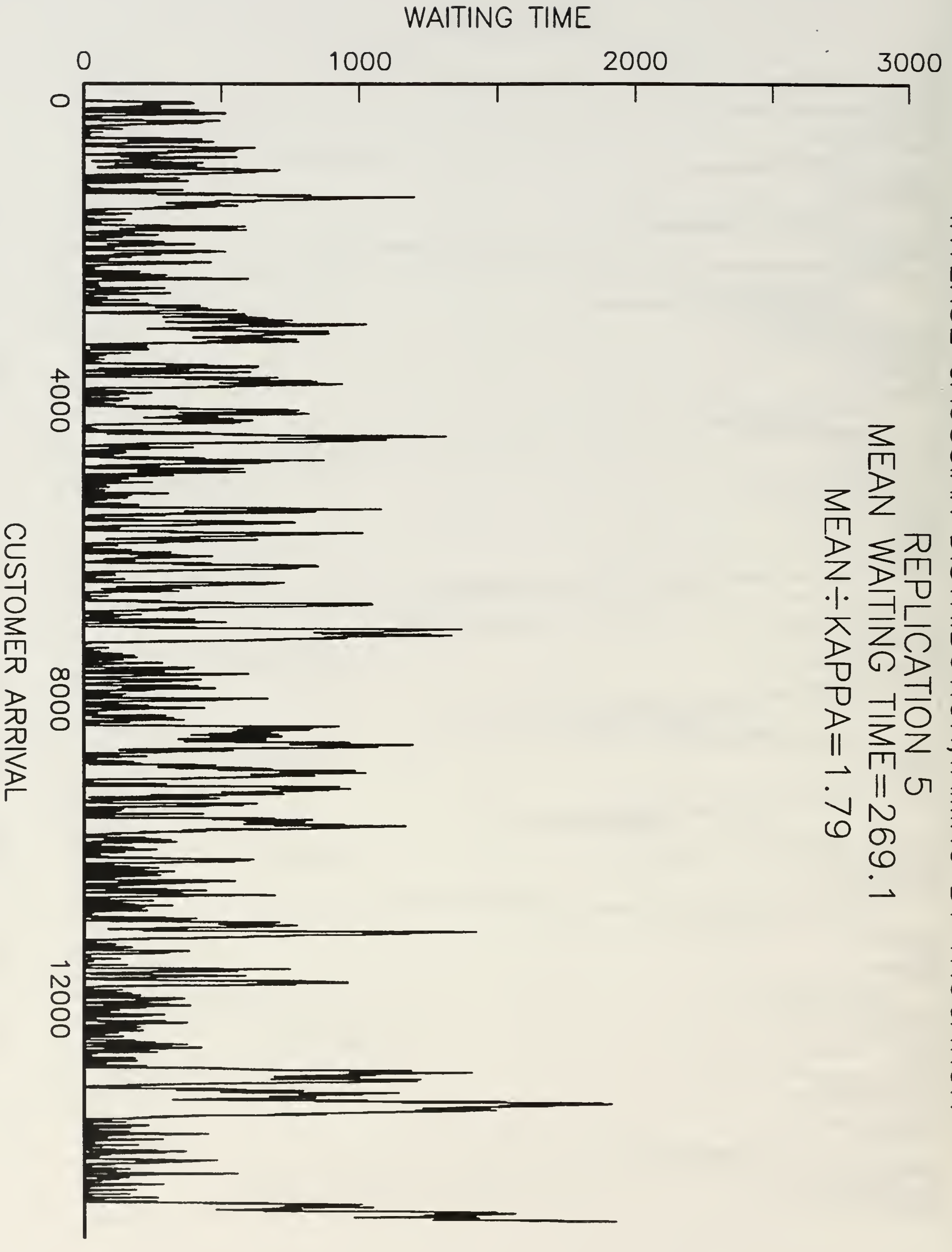

SUCCESSIVE CUSTOMER WAITING TIMES:KAPPA=150
INVERSE GAUSSIAN DISTRIBUTION;TAMING BY TRUCATION
REPLICATION 2
MEAN WAITING TIME=369.5
MEAN÷KAPPA=2.46
WAITING TIME
0
1000
2000
3000
0
4000
8000
12000
CUSTOMER ARRIVAL

Figure 1b



SUCCESSIVE CUSTOMER WAITING TIMES:KAPPA=150
INVERSE GAUSSIAN DISTRIBUTION;TAMING BY TRUCATION
REPLICATION 3
MEAN WAITING TIME=324.1
MEAN÷KAPPA=2.16

Figure 1c



SUCCESSIVE CUSTOMER WAITING TIMES:KAPPA=150
INVERSE GAUSSIAN DISTRIBUTION;TAMING BY TRUCATION
REPLICATION 5
MEAN WAITING TIME=269.1
MEAN÷KAPPA=1.79

Figure 1d

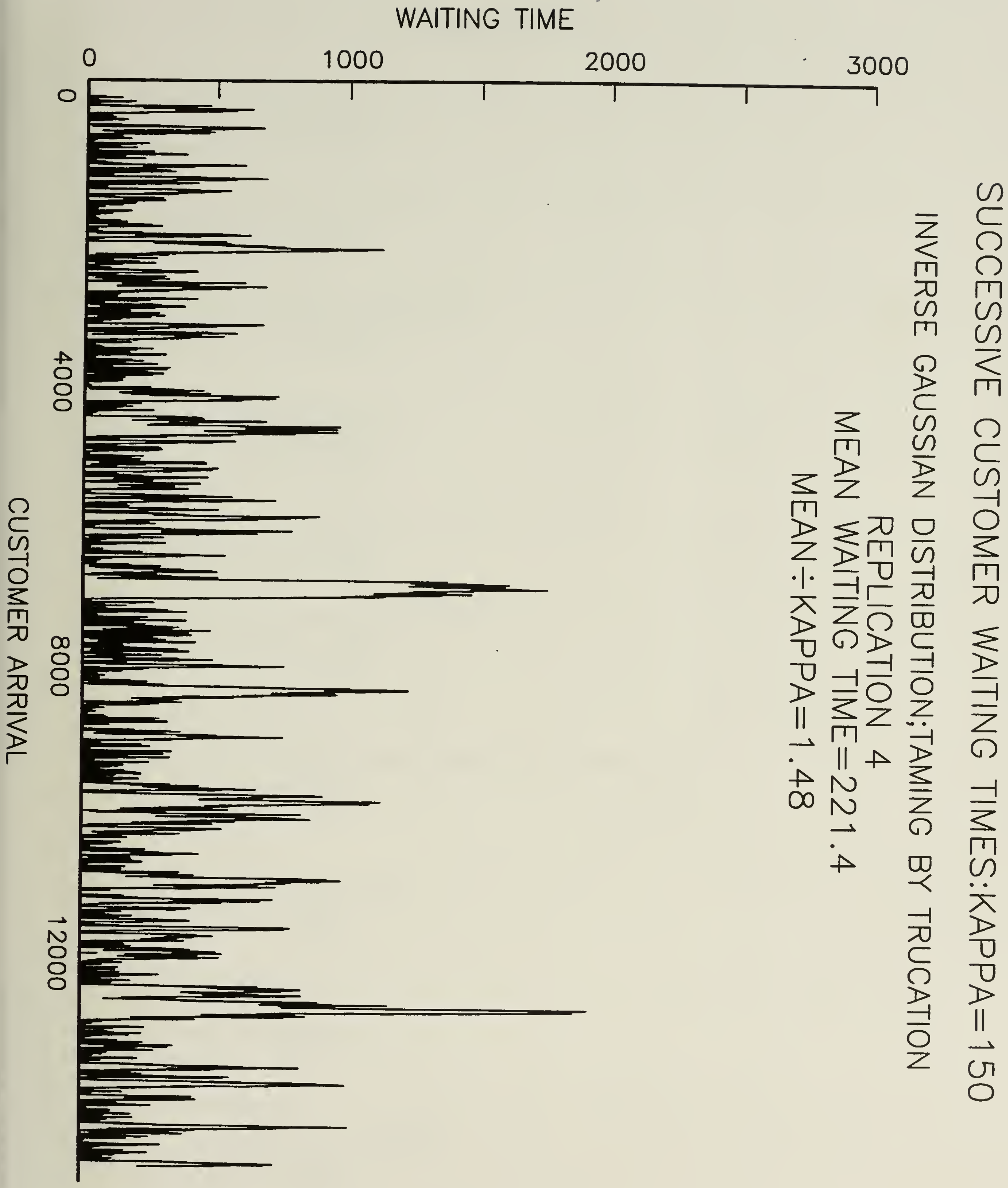

SUCCESSIVE CUSTOMER WAITING TIMES:KAPPA=150
INVERSE GAUSSIAN DISTRIBUTION;TAMING BY TRUCATION
REPLICATION 4
MEAN WAITING TIME=221.4
MEAN÷KAPPA=1.48
WAITING TIME
0
1000
2000
3000
CUSTOMER ARRIVAL
0
4000
8000
12000

Figure 1e

DISTRIBUTION LIST

1. Research Office (Code 09) 1
Naval Postgraduate School
Monterey, CA 93943-5000

2. Dudley Knox Library (Code 013) 2
Naval Postgraduate School
Monterey, CA 93943-5002

3. Defense Technical Information Center 2
8725 John J. Kingman Rd., STE 0944
Ft. Belvoir, VA 22060-6218

4. Therese Bilodeau 1
Dept of Operations Research
Naval Postgraduate School
Monterey, CA 93943-5000

5. Prof. Donald P. Gaver (Code OR/Gv) 5
Naval Postgraduate School
Monterey, CA 93943-5000

6. Prof. Patricia A. Jacobs (Code OR/Jc) 5
Naval Postgraduate School
Monterey, CA 93943-5000

7. Dr. J. Abrahams 1
Code 111, Room 607
Mathematical Sciences Division, Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

8. Prof. D. R. Barr 1
Dept. of Systems Engineering
U.S. Military Academy
West Point, NY 10996

9. Dr. David Brillinger 1
Statistics Dept.
University of California
Berkeley, CA 94720

10. Dr. David Burman 1
AT&T Bell Telephone Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

11. Center for Naval Analyses 1
4401 Ford Avenue
Alexandria, VA 22302-0268

12. Prof. H. Chernoff.................... 1
Department of Statistics
Harvard University
1 Oxford Street
Cambridge, MA 02138

13. Mr. Wm. P. Clay.................... 1
Director, USAMSAA
Attn: AMXSY-CA
APG, MD 21005-5071

14. Dr. Edward G. Coffman, Jr..................... 1
AT&T Bell Telephone Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

15. Prof. John Copas 1
Dept. of Statistics
University of Warwick
Coventry CV4 7AL
ENGLAND

16. Prof. Sir David Cox.................... 1
Nuffield College
Oxford OXI INF
ENGLAND

17. Prof. H. G. Daellenbach 1
Dept. of Operations Research
University of Canterbury
Christchurch
NEW ZEALAND

18. Dr. S. R. Dalal 1
Dept. of Statistics
University of Michigan
Ann Arbor, MI 48109

19. Dr. D. F. Daley.................... 1
Statistics Dept. (I.A.S.)
Australian National University
Canberra, A.C.T 2606
AUSTRALIA

20. Mr. Mike Davis 1
R55
9800 Savage Road
Ft. Meade, MD 20755

21. Dr. B. Doshi 1
AT&T Bell Laboratories
HO 3M-335
Holmdel, NJ 07733

22. Dr. Guy Fayolle 1
I.N.R.I.A.
Dom de Voluceau-Rocquencourt
78150 Le Chesnay Cedex
FRANCE

23. Dr. M. J. Fischer 1
Defense Communications Agency
1860 Wiehle Avenue
Reston, VA 22070

24. Prof. George S. Fishman 1
Curr. in OR & Systems Analysis
University of North Carolina
Chapel Hill, NC 20742

25. Dr. Neil Gerr 1
Office of Naval Research
Arlington, VA 22217

26. Dr. R. J. Gibbens 1
Statistics Laboratory
16 Mill Lane
Cambridge
ENGLAND

27. Prof. Peter Glynn 1
Dept. of Operations Research
Stanford University
Stanford, CA 94305

28. Prof. Linda V. Green 1
Graduate School of Business
Columbia University
New York, NY 10027

29. Dr. Shlomo Halfin...1
Bellcore
445 South Street
Morristown, NJ 07962-1910
(MRE 2L309)

30. Prof. Bernard Harris 1
Dept. of Statistics
University of Wisconsin
610 Walnut Street
Madison, WI 53706

31. Prof. J. Michael Harrison 1
Graduate School of Business
Stanford University
Stanford, CA 94305-5015

32. Dr. P. Heidelberger...1
IBM Research Laboratory
Yorktown Heights
New York, NY 10598

33. Prof. D. L. Iglehart 1
Dept. of Operations Research
Stanford University
Stanford, CA 94305-5015

34. Dr. Richard Ivanetich 1
Institute for Defense Analysis
1800 North Beauregard
Alexandria, VA 22311

35. Dr. F. P. Kelly 1
Statistics Laboratory
16 Mill Lane
Cambridge
ENGLAND

36. Dr. Jon Kettenring...1
Bellcore
445 South Street
Morris Township, NJ 07962-1910

37. Mr. Koh Peng Kong..1
OA Branch, DSO
Ministry of Defense
Blk 29 Middlesex Road
SINGAPORE 1024

38. Prof. Guy Latouche 1
University Libre Bruxelles
C.P. 212, Blvd. De Triomphe
Bruxelles B-1050
BELGIUM

39. Dr. A. J. Lawrance 1
Dept. of Mathematics
University of Birmingham
P.O. Box 363
Birmingham B15 2TT
ENGLAND

40. Prof. M. Leadbetter 1
Department of Statistics
University of North Carolina
Chapel Hill, NC 27514

41. Prof. J. Lehoczky 1
Department of Statistics
Carnegie-Mellon University
Pittsburgh, PA 15213

42. Dr. D. M. Lucantoni 1
AT&T Bell Laboratories
Holmdel, NJ 07733

43. Dr. James R. Maar 1
National Security Agency
9608 Basket Ring
Columbia, MD 21045-0689

44. Dr. Colin Mallows 1
AT&T Bell Telephone Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

45. Prof. M. Mazumdar 1
Dept. of Industrial Engineering
University of Pittsburgh
Pittsburgh, PA 15235

46. Dr. James McKenna 1
Bell Communications Research
445 South Street
Morristown, NJ 07960-1910

47. Prof. Paul Moose 1
C3I Academic Group
Naval Postgraduate School
Monterey, CA 93943-5000

48. Dr. John A. Morrison 1
AT&T Bell Telephone Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

49. Dr. T. J. Ott 1
Bellcore
445 South Street
Morristown, NJ 07962-1910
(MRE 2P388)

50. Dr. V. Ramaswami 1
MRE 2Q-358
Bell Communications Research Inc.
445 South Street
Morristown, NJ 07960

51. Dr. Martin Reiman 1
Rm #2C-117
AT&T Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974-2040

52. Prof. Maria Rieders 1
Dept. of Industrial Engineering
Northwestern University
Evanston, IL 60208

53. Dr. Rhonda Righter 1
Dept. of Decision & Info. Sciences
Santa Clara University
Santa Clara, CA 95118

54. Prof. Frank Samaniego 1
Statistics Department
University of California
Davis, CA 95616

55. Prof. G. A. F. Seber 1
Dept. of Statistics
Univ. of Auckland
Private Bag 92019
Auckland
NEW ZEALAND

56. Prof. G. Shanthikumar 1
The Management Science Group
School of Business Administration
University of California
Berkeley, CA 94720

57. Prof. N. D. Singpurwalla 1
George Washington University
Washington, DC 20052

58. Prof. H. Solomon 1
Department of Statistics
Sequoia Hall
Stanford University
Stanford, CA 94305

59. Prof. J. R. Thompson 1
Dept. of Mathematical Science
Rice University
Houston, TX 77001

60. Prof. J. W. Tukey 1
Statistics Dept., Fine Hall
Princeton University
Princeton, NJ 08540

61. Dr. D. Vere-Jones 1
Dept. of Math
Victoria Univ. of Wellington
P.O. Box 196
Wellington
NEW ZEALAND

62. Mr. Lai Kah Wah 1
Operations Analysis Department
Systems & Computer Organisation
19th Storey, Defence Technology Tower B
Depot Road 0410
SINGAPORE

63. Dr. Ed Wegman 1
George Mason University
Fairfax, VA 22030

64. Dr. L. Wein 1
Operations Research Center, Rm E40-164
Massachusetts Institute of Technology
Cambridge, MA 02139

65. Dr. Alan Weiss 1
Rm 2C-118
AT&T Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974-2040

66. Dr. P. Welch 1
IBM Research Laboratory
Yorktown Heights, NY 10598

67. Prof. Roy Welsch 1
Sloan School
M.I.T.
Cambridge, MA 02139